1 Abstract

THESIS: A Generalized Poisson Model for Gene Expression Profiling using RNA Sequence Data

STUDENT: **Paromita S Nitu**

DEGREE: Master of Science

COLLEGE: Sciences and Humanities

DATE: May, 2016

PAGES: 42

An RNA-Seq experiment is a useful tool in characterizing and quantifying transcriptomes into read counts and identifying differentially expressed (DE) genes under different treatment conditions. However, analyzing RNA-Seq data in the quest of differentially expressed genes is not straight forward. Focusing on the experiment of interest, different approaches can be made in identifying DE genes. Here, we propose to use a two parameter generalized poisson (GP) model to address the non-uniformity of read counts than the traditional poisson model and apply it to Arabidopsis pilot survey data by TCC(http://bioconductor .org/packages/release/bioc/html/TCC.html). A comparison study has also been performed with built in R-packages edgeR and DESeq with their default settings to understand the performance of GP model. Here, 28 new differentially expressed genes have been identified by GP model more than edgeR and DESeq for Arabidopsis data and these genes can be a potential source of information in treating bacterial infec-

tion to the experimenters. Therefore, the approach of using GP model in real data set evident a significant performance to the in built methods of R-packages.